APPENDIX C

second edition

CMOS DIGITAL INTEGRATED CIRCUITS

Analysis and Design

SUNG-MO (STEVE) KANG
University of Illinois at Urbana- Champaign

YUSUF LEBLEBICI
Worcester Polytechnic Institute
Swiss Federal Institute of Technology-Lausanne

WCB McGraw-Hill

Boston Burr Ridge, IL Dubuque, IA Madison, WI New York San Francisco St. Louis
Bangkok Bogotá Caracas Lisbon London Madrid
Mexico City Milan New Delhi Seoul Singapore Sydney Taipei Toronto

WCB/McGraw-Hill

A Division of The McGraw-Hill Companies

To Myoung-A, Jennifer and Jefferey,
and Anil and Ebru

CMOS DIGITAL INTEGRATED CIRCUITS: ANALYSIS AND DESIGN

This book is printed on acid-free paper.

1 2 3 4 5 6 7 8 9 0 DOC/DOC 9 3 2 1 0 9 8

ISBN 0-07-292507-8

Vice president and editorial director: Kevin T. Kane
Publisher: Thomas Casson
Executive editor: Elizabeth A. Jones
Senior developmental editor: Kelley Butcher
Marketing manager: John T. Wannemacher
Project manager: Karen Nelson
Senior production supervisor: Heather D. Burbridge
Senior designer: Lorna Lo
Cover designer: Nicole Leong
Supplement coordinator: Linda Huenecke
Typeface: 10/12 Times Roman
Printer: R. R. Donnelley & Sons Company

Library of Congress Cataloging-in-Publication Data

Kang, Sung Mo, (date)
CMOS digital integrated circuits: analysis and design/Sung-Mo (Steve) Kang, Yusuf Leblebici.
–2nd ed.
p. cm.
Includes index.
ISBN 0-07-292507-8
1. Metal oxide semiconductors, Complementary. 2. Digital integrated circuits. I. Leblebici, Yusuf.
II. Title.
TK7871.99M44K36 1999
621.39'5--dc21

98-28075

http://www.mhhe.com

$$x_{dm} = \sqrt{\frac{2 \cdot \varepsilon_{Si} \cdot |2\phi_F|}{q \cdot N_A}} \tag{3.13}$$

The creation of a conducting surface inversion layer through externally applied gate bias is an essential phenomenon for current conduction in MOS transistors. In the following section, we will examine the structure and the operation of the MOS Field Effect Transistor (MOSFET).

3.3. Structure and Operation of MOS Transistor (MOSFET)

The basic structure of an n-channel MOSFET is shown in Fig. 3.8. This four-terminal device consists of a p-type substrate, in which two n^+ diffusion regions, the drain and the source, are formed. The surface of the substrate region between the drain and the source is covered with a thin oxide layer, and the metal (or polysilicon) gate is deposited on top of this gate dielectric. The midsection of the device can easily be recognized as the basic MOS structure which was examined in the previous sections. The two n^+ regions will be the current-conducting terminals of this device. Note that the device structure is completely symmetrical with respect to the drain and source regions; the different roles of these two regions will be defined only in conjunction with the applied terminal voltages and the direction of the current flow.

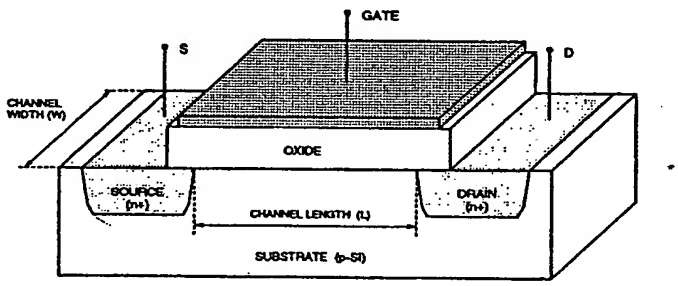

Figure 3.8. The physical structure of an n-channel enhancement-type MOSFET.

A conducting *channel* will eventually be formed through applied gate voltage in the section of the device between the drain and the source diffusion regions. The distance between the drain and source diffusion regions is the *channel length L*, and the lateral extent of the channel (perpendicular to the length dimension) is the *channel width W*. Both the channel length and the channel width are important parameters which can be used to control some of the electrical properties of the MOSFET. The thickness of the oxide layer covering the channel region, t_{ox}, is also an important parameter.

A MOS transistor which has no conducting channel region at zero gate bias is called an *enhancement-type* (or *enhancement-mode*) MOSFET. If a conducting channel already exists at zero gate bias, on the other hand, the device is called a *depletion-type* (or

depletion-mode) MOSFET. In a MOSFET with p-type substrate and with n^+ source and drain regions, the channel region to be formed on the surface is n-type. Thus, such a device with p-type substrate is called an *n-channel MOSFET*. In a MOSFET with n-type substrate and with p^+ source and drain regions, on the other hand, the channel is p-type and the device is called a *p-channel MOSFET*.

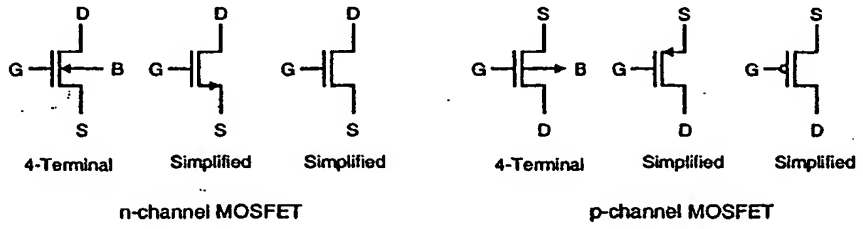

Figure 3.9. Circuit symbols for n-channel and p-channel enhancement-type MOSFETs.

The abbreviations used for the device terminals are: G for the gate, D for the drain, S for the source, and B for the substrate (or body). In an n-channel MOSFET, the source is defined as the n^+ region which has a *lower* potential than the other n^+ region, the drain. By convention, all terminal voltages of the device are defined with respect to the source potential. Thus, the gate-to-source voltage is denoted by V_{GS}, the drain-to-source voltage is denoted by V_{DS}, and the substrate-to-source voltage is denoted by V_{BS}. Circuit symbols for both n-channel and p-channel enhancement-type MOSFETs are shown in Fig. 3.9. While the four-terminal symbolic representation shows all external terminals of the device, the simple three-terminal representation will also be used extensively. Note that in the simple MOSFET circuit symbol, the small arrow always marks the source terminal.

Consider first the n-channel enhancement-type MOSFET shown in Fig. 3.8. The simple operation principle of this device is: *control the current conduction between the source and the drain, using the electric field generated by the gate voltage as a control variable*. Since the current flow in the channel is also controlled by the drain-to-source voltage and by the substrate voltage, the current can be considered a function of these external terminal voltages. We will examine in detail the functional relationships between the channel current (also called the *drain current*) and the terminal voltages. In order to start current flow between the source and the drain regions, however, we have to form a conducting channel first.

The simplest bias condition that can be applied to the n-channel enhancement-type MOSFET is shown in Fig. 3.10. The source, the drain, and the substrate terminals are all connected to ground. A positive gate-to-source voltage V_{GS} is then applied to the gate in order to create the conducting channel underneath the gate. With this bias arrangement, the channel region between the source and the drain diffusions behaves exactly the same as for the simple MOS structure we examined in Section 3.2. For small gate voltage levels, the majority carriers (holes) are repelled back into the substrate, and the surface of the p-type substrate is depleted. Since the surface is devoid of any mobile carriers, current conduction between the source and the drain is not possible.

by selective dopant ion implantation into the channel region of the MOSFET. For n-channel MOSFETs, the threshold voltage is *increased* (made more positive) by adding extra p-type impurities (acceptor ions). Alternatively, the threshold voltage of th n-channel MOSFET can be *decreased* (made more negative) by implanting n-type impurities (dopant ions) into the channel region.

The amount of change in the threshold voltage as a result of extra implants can be approximated as follows. Let the density of implanted impurities be represented by N_I [cm^{-2}]. Assume that all implanted ions are electrically active, i.e., each ion contributes to the depletion region charge. Then, the threshold voltage V_{T0} at zero substrate bias ($V_{SB} = 0$) will be shifted by an amount of qN_I/C_{ox}. This approximation obviously neglects the variation of the substrate Fermi level ϕ_F as the result of extra implants, but it nevertheless provides a fair estimate for the threshold voltage shift.

Exercise 3.1

Consider the following p-channel MOSFET process:

Substrate doping $N_D = 10^{15}$ cm^{-3}, polysilicon gate doping density $N_D = 10^{20}$ cm^{-3}, gate oxide thickness $t_{ox} = 650$ Å, and oxide-interface charge density $N_{ox} = 2 \times 10^{10}$ cm^{-2}. Use $\varepsilon_{Si} = 11.7\varepsilon_0$ and $\varepsilon_{ox} = 3.97\varepsilon_0$ for the dielectric coefficients of silicon and silicon-dioxide, respectively.

(a) Calculate the threshold voltage V_{T0}, for $V_{SB} = 0$.

(b) Determine the type and the amount of channel ion implantation which are necessary to achieve a threshold voltage of $V_{T0} = -2$ V.

Note that, using selective ion implantation into the channel, the threshold voltage of an n-channel MOSFET can also be made negative. This means that the resulting nMOS transistor will have a conducting channel at $V_{GS} = 0$, enabling current flow between its source and drain terminals as long as V_{GS} is larger than the negative threshold voltage. Such a device is called a *depletion-type* (or *normally-on*) n-channel MOSFET. We will see several practical applications for depletion-type nMOS transistors in the design of MOS digital circuits. Except for its negative threshold voltage, the depletion-type n-channel MOSFET exhibits the same electrical behavior as the enhancement-type n-channel MOSFET. Figure 3.13 shows the conventional circuit symbols used for depletion-type n-channel MOSFETs.

source voltages, at which the lateral electric field in the drain end of the channel accelerates the electrons. The electrons arriving at the Si-SiO_2 interface with enough kinetic energy to surmount the surface potential barrier are injected into the oxide. Electrons and holes generated by impact ionization also contribute to the charge injection. Note that the channel hot-electron current and the subsequent damage in the gate oxide are localized near the drain junction (Fig. 3.27).

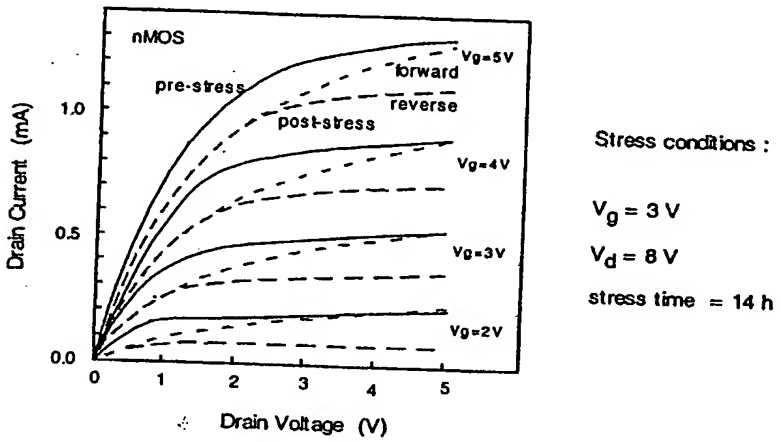

Figure 3.28. Typical drain current vs. drain voltage characteristics of an n-channel MOS transistor before and after hot-carrier induced oxide damage.

The hot-carrier induced damage in nMOS transistors has been found to result in either trapping of carriers on defect sites in the oxide or the creation of interface states at the silicon-oxide interface, or both. The damage caused by hot-carrier injection affects the transistor characteristics by causing a degradation in transconductance, a shift in the threshold voltage, and a general decrease in the drain current capability (Fig. 3.28). This performance degradation in the devices leads to the degradation of circuit performance over time. Hence, new MOSFET technologies based on smaller device dimensions must carefully account for the hot-carrier effects and also ensure reliable long-term operation of the devices.

Other reliability concerns for small-geometry devices include interconnect damage through electromigration, electrostatic discharge (ESD) and electrical over-stress (EOS).

3.6. MOSFET Capacitances

The majority of the topics covered in this chapter has been related to the steady-state behavior of the MOS transistor. The current-voltage characteristics investigated here can be applied for investigating the DC response of MOS circuits under various operating conditions. In order to examine the transient (AC) response of MOSFETs and digital

circuits consisting of MOSFETs, on the other hand, we have to determine the nature and the amount of parasitic capacitances associated with the MOS transistor.

The on-chip capacitances found in MOS circuits are in general complicated functions of the layout geometries and the manufacturing processes. Most of these capacitances are not lumped, but *distributed*, and their exact calculations would usually require complex, three-dimensional nonlinear charge-voltage models. In the following, we will develop simple approximations for the on-chip MOSFET capacitances that can be used in most hand calculations. These capacitance models are sufficiently accurate to represent the crucial characteristics of MOSFET charge-voltage behavior, and the equations are all based on fundamental semiconductor device theory, which should be familiar to most readers. We will also stress the distinction between the device-related capacitances and the interconnect capacitances. The capacitive contribution of metal interconnections between various devices is a very important component of the total parasitic capacitance observed in digital circuits. The estimation of this interconnect capacitance will be handled in Chapter 6.

Figure 3.29 shows the cross-sectional view and the top view (mask view) of a typical n-channel MOSFET. Until now, we concentrated on the cross-sectional view of the device, since we were primarily concerned with the flow of carriers within the MOSFET. As we study the parasitic device capacitances, we will have to become more familiar with the top view of the MOSFET. In this figure, the *mask length* (drawn length) of the gate is indicated by L_M, and the actual channel length is indicated by L. The extent of both the gate-source and the gate-drain overlap are L_D; thus, the channel length is given by

$$L = L_M - 2 \cdot L_D \tag{3.93}$$

Note that the source and drain overlap region lengths are usually equal to each other because of the symmetry of the MOSFET structure. Typically, L_D is on the order of 0.1 μm. Both the source and the drain diffusion regions have a width of W. The typical diffusion region length is denoted by Y. Note that both the source diffusion region and the drain diffusion region are surrounded by a p^+ doped region, also called the channel-stop implant. As the name indicates, the purpose of this additional p^+ region is to prevent the formation of any unwanted (parasitic) channels between two neighboring n^+ diffusion regions, i.e., to ensure that the surface between two such regions cannot be inverted. Hence, the p^+ channel-stop implants act to electrically isolate neighboring devices built on the same substrate.

We will identify the parasitic capacitances associated with this typical MOSFET structure as lumped equivalent capacitances *observed* between the device terminals (Fig. 3.30), since such a lumped representation can be easily used to analyze the dynamic transient behavior of the device. The reader must always be reminded, however, that in reality most parasitic device capacitances are due to three-dimensional, distributed charge-voltage relations within the device structure. Based on their physical origins, the parasitic device capacitances can be classified into two major groups: *oxide-related capacitances* and *junction capacitances*. First, the oxide-related capacitances will be considered.

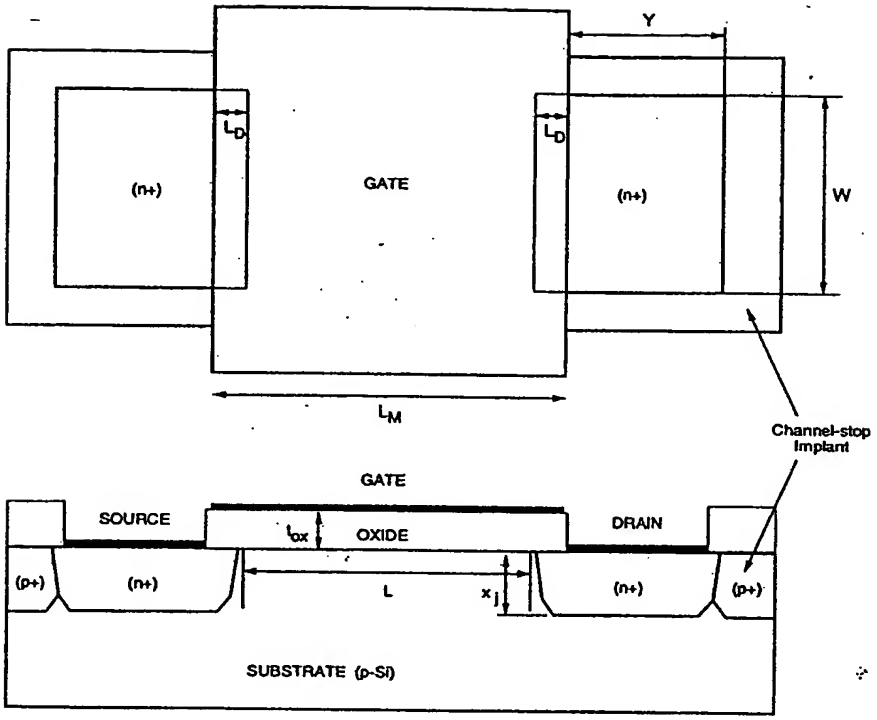

Figure 3.29. Cross-sectional view and top view (mask view) of a typical n-channel MOSFET.

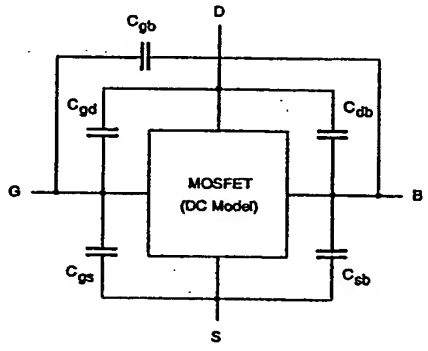

Figure 3.30. Lumped representation of the parasitic MOSFET capacitances.

It was shown earlier that the gate electrode overlaps both the source region and the drain region at the edges. The two overlap capacitances that arise as a result of this structural arrangement are called C_{GD} *(overlap)* and C_{GS} *(overlap)*, respectively. Assuming that both the source and the drain diffusion regions have the same width W, the overlap capacitances can be found as

$$\begin{aligned} C_{GS}(overlap) &= C_{ox} \cdot W \cdot L_D \\ C_{GD}(overlap) &= C_{ox} \cdot W \cdot L_D \end{aligned} \tag{3.94}$$

with

$$C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}} \tag{3.95}$$

Note that both of these overlap capacitances do not depend on the bias conditions, i.e., they are voltage-independent.

Now consider the capacitances which result from the interaction between the gate voltage and the channel charge. Since the channel region is connected to the source, the drain, and the substrate, we can identify three capacitances between the gate and these regions, i.e., C_{gs}, C_{gd}, and C_{gb}, respectively. Notice that in reality, the gate-to-channel capacitance is distributed and voltage-dependent. Then, the gate-to-source capacitance C_{gs} is actually the gate-to-channel capacitance *seen* between the gate and the source terminals; the gate-to-drain capacitance C_{gd} is actually the gate-to-channel capacitance *seen* between the gate and the drain terminals. A simplified view of their bias-dependence can be obtained by observing the conditions in the channel region during cut-off, linear, and saturation modes.

In cut-off mode (Fig. 3.31(a)), the surface is not inverted. Consequently, there is no conducting channel that links the surface to the source and to the drain. Therefore, the gate-to-source and the gate-to-drain capacitances are both equal to zero: $C_{gs} = C_{gd} = 0$. The gate-to-substrate capacitance can be approximated by

$$C_{gb} = C_{ox} \cdot W \cdot L \tag{3.96}$$

In linear-mode operation, the inverted channel extends across the MOSFET, between the source and the drain (Fig. 3.31(b)). This conducting inversion layer on the surface effectively shields the substrate from the gate electric field; thus, $C_{gb} = 0$. In this case, the distributed gate-to-channel capacitance may be viewed as being shared equally between the source and the drain, yielding

$$C_{gs} \equiv C_{gd} \equiv \frac{1}{2} \cdot C_{ox} \cdot W \cdot L \tag{3.97}$$

When the MOSFET is operating in saturation mode, the inversion layer on the surface does not extend to the drain, but it is pinched off (Fig. 3.31(c)). The gate-to-drain

capacitance component is therefore equal to zero ($C_{gd} = 0$). Since the source is still linked to the conducting channel, its shielding effect also forces the gate-to-substrate capacitance to be zero, $C_{gb} = 0$. Finally, the distributed gate-to-channel capacitance as seen between the gate and the source can be approximated by

$$C_{gs} \cong \frac{2}{3} \cdot C_{ox} \cdot W \cdot L \tag{3.98}$$

Table 3.6 lists a summary of the approximate oxide capacitance values in three different operating modes of the MOSFET. The variation of the distributed parasitic oxide capacitances as functions of the gate-to-source voltage V_{GS} is also shown in Fig. 3.32.

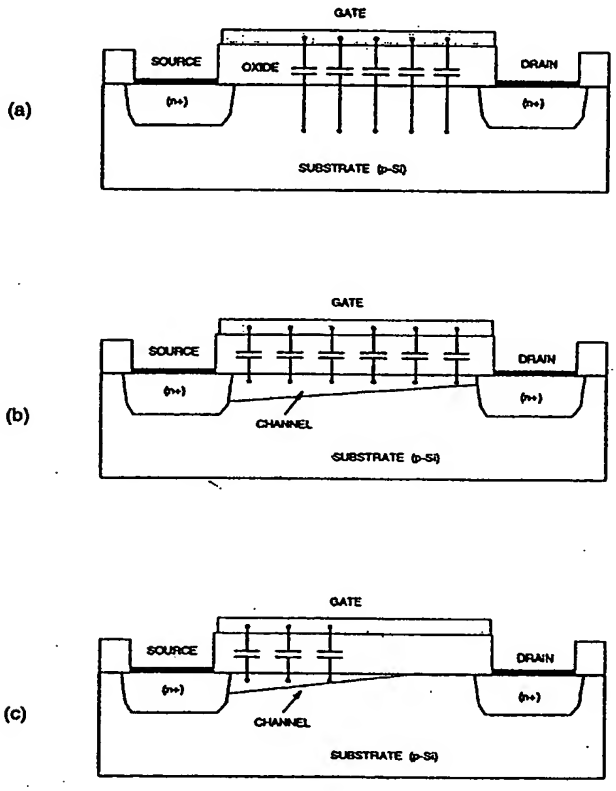

Figure 3.31. Schematic representation of MOSFET oxide capacitances during (a) cut-off, (b) linear, and (c) saturation modes.

Obviously, we have to combine the distributed C_{gs} and C_{gd} values found here with the relevant verlap capacitance values, in order to calculate the total capacitance between the external device terminals. It is also worth mentioning that the sum of all three voltage-dependent (distributed) gate oxide capacitances ($C_{gb} + C_{gs} + C_{gd}$) has a minimum value of $0.66\,C_{ox}WL$ (in saturation mode) and a maximum value of $C_{ox}WL$ (in cut-off and linear modes). For simple hand calculations where all three capacitances can be considered to be connected in parallel, a constant worst-case value of $C_{ox}W(L+2L_D)$ can be used for the sum of MOSFET gate oxide capacitances.

Capacitance	Cut-off	Linear	Saturation
C_{gb} (total)	$C_{ox}WL$	0	0
C_{gd} (total)	$C_{ox}WL_D$	$\frac{1}{2}C_{ox}WL + C_{ox}WL_D$	$C_{ox}WL_D$
C_{gs} (total)	$C_{ox}WL_D$	$\frac{1}{2}C_{ox}WL + C_{ox}WL_D$	$\frac{2}{3}C_{ox}WL + C_{ox}WL_D$

Table 3.6. Approximate oxide capacitance values for three operating modes of the MOS transistor.

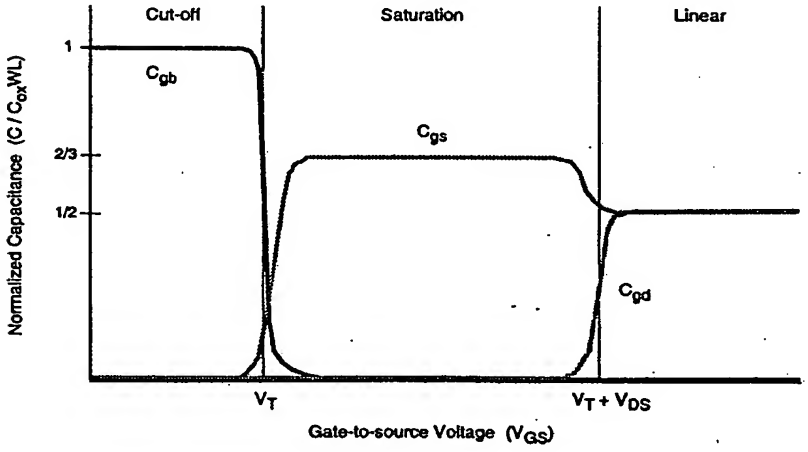

Figure 3.32. Variation of the distributed (gate-to-channel) oxide capacitances as functions of gate-to-source voltage V_{GS}.